

**PCT**

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification 5 :</b> <b>C12N 15/13, C12P 21/08</b>		<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 93/06217</b> <b>(43) International Publication Date:</b> 1 April 1993 (01.04.93)
<b>(21) International Application Number:</b> PCT/US92/07986			<b>(72) Inventor; and</b> <b>(75) Inventor/Applicant (for US only):</b> CARTER, Paul, J. [GB/US]; 2074 18th Avenue, San Francisco, CA 94116 (US).
<b>(22) International Filing Date:</b> 18 September 1992 (18.09.92)			<b>(74) Agents:</b> ADLER, Carolyn, R. et al.; Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080-4990 (US).
<b>(30) Priority data:</b> 07/762,292	19 September 1991 (19.09.91)	US	<b>(81) Designated States:</b> CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, SE).
<b>(60) Parent Application or Grant</b> <b>(63) Related by Continuation</b> US	07/762,292 (CIP)	Filed on	<b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
19 September 1991 (19.09.91)			
<b>(71) Applicant (for all designated States except US):</b> GENENTECH, INC. [US/US]; 460 Point San Bruno Boulevard, South San Francisco, CA 94080-4990 (US).			

**(54) Title:** EXPRESSION IN E. COLI OF ANTIBODY FRAGMENTS HAVING AT LEAST A CYSTEINE PRESENT AS A FREE THIOL, USE FOR THE PRODUCTION OF BIFUNCTIONAL F(ab')<sub>2</sub> ANTIBODIES

**(57) Abstract**

Methods for the high yield production of antibody Fv-containing polypeptides, especially Fab' and F(ab')<sub>2</sub> antibody fragments are provided. Expression of heavy and light chain Fv in a microbial secretory system is followed by recovery of Fv from the periplasm under conditions that maintain a cysteine residue as a free thiol. The free thiol is reacted with free thiol of an antibody fragment of the same or differing specificity, or with agents such as diagnostic labels or therapeutic moieties. The products offer advantages of homogeneity and purity not available through the use of known methods for preparing such derivatives.

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	MN	Mongolia
AU	Australia	FR	France	MR	Mauritania
BB	Barbados	GA	Gabon	MW	Malawi
BE	Belgium	GB	United Kingdom	NL	Netherlands
BF	Burkina Faso	GN	Guinea	NO	Norway
BG	Bulgaria	GR	Greece	NZ	New Zealand
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IE	Ireland	PT	Portugal
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	RU	Russian Federation
CG	Congo	KP	Democratic People's Republic of Korea	SD	Sudan
CH	Switzerland	KR	Republic of Korea	SE	Sweden
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovak Republic
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CS	Czechoslovakia	LU	Luxembourg	SU	Soviet Union
CZ	Czech Republic	MC	Monaco	TD	Chad
DE	Germany	MG	Madagascar	TG	Togo
DK	Denmark	ML	Mali	UA	Ukraine
ES	Spain			US	United States of America

EXPRESSION IN E.COLI OF ANTIBODY FRAGMENTS HAVING AT LEAST A CYSTEINE PRESENT AS A FREE THIOL. USE FOR THE PRODUCTION OF BIFUNCTIONAL F(ab')<sub>2</sub> ANTIBODIES

5

Field of the Invention

This invention relates to the production of functional antibody fragments in a microbial host.

10

Background of the Invention

Naturally occurring antibodies (immunoglobulins) comprise two heavy chains linked together by disulfide bonds and two light chains, each light chain being linked 15 to one of the heavy chains by disulfide bonds. Each chain has an N-terminal variable domain ( $V_H$  or  $V_L$ ) and a constant domain at its C-terminus; the constant domain of the light chain is aligned with and disulfide bonded to the first constant domain of the heavy chain, and the light chain variable domain is aligned with the variable domain of the heavy chain. The heavy chain constant region includes (in the N- to C-terminal direction) the  $C_H1$  and hinge regions. The light chain also contains a hinge 20 domain. Particular amino acid residues are believed to form an interface between and disulfide bond the light and heavy chain variable domains, see e.g. Chothia *et al.*, *J. Mol. Biol.* 186:651-663 (1985); Novotny and Haber, *Proc. Natl. Acad. Sci. USA* 82:4592-4596 (1985); Padlan *et al.*, *Mol. Immunol.*, 23(9): 951-960 (1986); and S. 25 Miller, *J. Mol. Biol.*, 216: 965-973 (1990).

The constant domains are not involved directly in binding the antibody to an antigen, but are involved in various effector functions, such as participation of the antibody in antibody-dependent cellular cytotoxicity and complement dependent cytotoxicity. The variable domains of each pair of light and heavy chains are involved 30 directly in binding the antibody to the antigen. The domains of natural light and heavy chains have the same general structure, the so-called immunoglobulin fold, and each domain comprises four framework (FR) regions, whose sequences are somewhat conserved, connected by three hyper-variable or complementarity determining regions (CDRs) (see Kabat, E. A. *et al.*, *Sequences of Proteins of Immunological Interest*, 35 National Institutes of Health, Bethesda, MD, (1987)). The four framework regions

largely adopt a  $\beta$ -sheet conformation and the CDRs form loops connecting, and in some cases forming part of, the  $\beta$ -sheet structure. The CDRs in each chain are held in close proximity by the framework regions and, with the CDRs from the other chain, contribute to the formation of the antigen binding site.

5       Antibodies can be divided into a variety of antigen-binding fragments. The  $F_v$  fragment is a heterodimer containing only the variable domains of the heavy chain and the light chain. The Fab fragment also contains the constant domain of the light chain and the first constant domain ( $C_H1$ ) of the heavy chain. Fab' fragments differ from Fab fragments by the addition of a few residues at the carboxy terminus of the 10 heavy chain  $C_H1$  domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group.  $F(ab')_2$  antibody fragments originally were produced as pairs of Fab' fragments which are between hinge cysteines.

15       Recombinant production of antibodies and antibody fragments facilitates the engineering of antibodies with enhanced antigen binding affinities, modified immunogenicity, and also of bifunctional antibodies. The first expression systems reported with which functional antibodies could be obtained were for mammalian cells. The U.S. patent of Cabilly *et al.*, U.S. patent No. 4,816,567, for example, teaches a method of co-expressing at least the variable region sequences of light and heavy 20 chains in host cells. Other researchers in the field have reported baculovirus expression systems (Haseman *et al.*, *Proc. Natl. Acad. Sci. USA* 87:3942-3946 (1990), yeast systems (Horwitz *et al.*, *Proc. Natl. Acad. Sci. USA*, 85:8678-8682 1988), combinatorial libraries in phage lambda (Huse *et al.*, *Science* 246:1275-1281 (1989), and work with filamentous phage (McCafferty *et al.*, *Nature* 348:552-554 25 (1990).

25       The production of antibodies and antibody fragments in bacterial systems have been pursued by workers in the field, particularly in *E. coli* expression systems. There are several advantages to *E. coli* expression systems, including a well-studied and convenient gene technology which permits constructs to be made easily and directly expressed, and the relatively convenient and economical large-scale production of product made possible by the fast growth of *E. coli* and its comparatively simple fermentation. The large-scale production of functional antibody fragments in *E. coli* would be valuable for research as well as commercial applications.

30       The expression of antibody genes in bacteria was reported by Cabilly *et al.*, *Proc. Natl. Acad. Sci. USA* 81:3273-3277 (1984), Boss *et al.*, *Nucleic Acids Res.*

-3-

12:3791-3806 (1984); these reports show cytoplasmic expression and rather variable yields were reported. Zemel-Dreasen *et al.*, *Gene* 315-322 (1984) report the secretion and processing of an immunoglobulin light chain in *E. coli*. Plückthun *et al.*, *Cold Spring Harbor Symposia on Quantitative Biology*, Volume LII, pages 105-112 (1987, Cold Spring Harbor Laboratory) disclose expression of a cytoplasmic hybrid protein, a potentially exportable hybrid protein, and expression and periplasmic transport of  $V_L$ ,  $V_H$ ,  $V_L C_L$ , and  $V_H C_H$  chains as fusions with an alkaline phosphatase or  $\beta$ -lactamase signal sequences. Skerra and Plückthun, *Science* 240:1038-1041 (1988) report the periplasmic secretion and correct folding *in vivo* of the variable domains of an antibody to the *E. coli* periplasm; a similar strategy and results were reported by Better *et al.*, *Science* 240:1041-1043 (1988) for expression of a murine Fab fragment.

Bird *et al.*, *Nature* 332:323-327 (1988) report the linkage of the light and heavy chain fragment of the Fv region *via* an amino acid sequence, and production of the complex as a single polypeptide in *E. coli*; see also Ladner *et al.*, U.S. patent 4,946,778. Huston *et al.*, *Proc. Natl. Acad. Sci. USA* 85:5879-5883 (1988) report similar work. Ward *et al.*, *Nature* 341:544-546 (1989) report the production in and secretion of "single-domain" antibodies (isolated heavy chain variable domains) from *E. coli*. Condra *et al.*, *Journal of Biological Chemistry*, 265(4):2292-2295 (1990) disclose the expression of cDNAs encoding antibody light and heavy chains in *E. coli* and their renaturation into Fab fragments. Better and Horwitz, *Methods in Enzymology*, 178:476-496 (1989), describe the expression, and secretion of functional Fab fragments from *E. coli* and *Saccharomyces cerevisiae*.

Plückthun and Skerra describe techniques for the expression of functional antibody Fv and Fab fragments in *E. coli* in *Methods In Enzymology* 178:497-515 (1989). According to their strategy, in the cytoplasm, the precursor proteins for  $V_L$  and  $V_H$ , each fused to a bacterial signal sequence, are synthesized in reduced form. After translocation through the inner membrane into the periplasm, the signal sequences are cleaved, the domains fold and assemble, and the disulfide bonds form. They teach that expression of the Fab fragment according to their strategy is analogous. Similar expression strategies are found elsewhere in the literature. See also Plückthun, *Biotechnology*, 9:545-551 (1991) for a review of *E. coli* expression of antibody fragments.

Cabilly (*Gene*, 85:553-557 [1989]) teaches that, in *E. coli* cells growing at reduced temperatures (21 °C or 30 °C, rather than at 37 °C), a single expression

-4-

plasmid coding for kappa-chains and truncated heavy chains (Fd fragments) gives rise to high yields of functional Fab fragments. Cabilly discusses that the Fab fragments seem to exist in the *E. coli* cytoplasm as non-covalently linked dimers, but that soluble Fab fragments isolated from *E. coli* appear as covalent dimers, formed by air 5 oxidation following cell rupture.

It is known in the literature that in the presence of low concentrations of a mild reductant such as cysteamine the bivalent  $F(ab')_2$  antibody fragment dissociates into two Fab' fragments. This dissociation is reversible by mild oxidation. The production 10 of Fab and  $F(ab')_2$  antibody fragments has also been shown by partial reduction and limited proteolysis of intact antibodies, see e.g. Parham, in *Cellular Immunology* (E.M. Weir, Ed., Blackwell Scientific, CA) 4th edition, vol. 1 chapter 14 (1983), however with these methods it is difficult to control the precise nature and proportions of the antibody fragment recovered. Bivalent antibodies are those which contain at least two epitopic combining sites (which sites may be on the same or different antigens).

15 Bispecific antibodies are bivalent antibodies capable of binding two epitopes not shared by a single antigen. Bispecific monoclonal antibodies (BsMAbs) with dual specificities for tumor-associated antigens on tumor cells and for surface markers in immune effector cells have been described (see, e.g. Liu *et al.*, *Proc. Natl. Acad. Sci. USA* 82:8648 (1985); Perez *et al.*, *Nature* 316:354 (1985)). These BsMAbs have 20 been shown to be effective in directing and triggering effector cells to kill tumor cell targets (Fanger *et al.*, *Immunol. Today* 12:51 (1991)). One approach to the production of BsMabs involves the fusion of two monoclonal antibody-producing hybridomas to form quadromas (hybrid hybridomas) which secrete BsMab in addition to undesirable chain combinations including parental MAbs (Milstein, C. and Cuello, 25 A. C., *Nature*, 305:537 [1983]). However, for production of bispecific *humanized* antibodies and antibody fragments, other techniques would be preferred.

Nisonoff and Mandy (*Nature* 4826:355-359 (1962)) describe the digestion of 30 rabbit antibodies and subsequent recombination of the antibody fragments; they disclose that antibody molecules of dual specificity can be obtained by combining univalent fragments of pepsin-treated antibodies of different specificities. See also Hammerling *et al.*, *Journal of Experimental Medicine* 128:1461-1469 (1968); Parham, *Human Immunology* 12:213-331 (1985); Raso and Griffin, *Cancer Research* 41:2073-2076 (1981); and Paulus (U.S. Patent 4,444,878).

Another approach utilizes directed chemical coupling of bispecific Fab' 35 fragments from two different MAbs to assemble a BsMAb, in this case a  $F(ab')_2$ , with

-5-

the desired specificities (See e.g., Nolan *et al.*, *Biochimica et Biophysica Acta* 1040:1 (1990). See also R. A. Maurer's Ph.D. Thesis, Harvard University (1978), and Brennan *et al.*, *Science* 229:81-83 (1985) for chemistries for the directed coupling of dithionitrobenzoate derivatives of Fab' fragments. Brennan *et al.* also teach the 5 use of use sodium arsenite to cross-link two proximate cysteines, however this reaction involves highly toxic compounds. (Glennie *et al.*, *J. Biol. Chem.*, 141(10): 3662-3670 [1985] and *J. Immunol.*, 139:2367-2375 [1975]) teach the preparation of bispecific F(ab')<sub>2</sub> antibody fragments containing thioether linkages. These chemistries would also be applicable for the coupling of identical Fab' fragments.

10 Lyons *et al.*, *Protein Engineering* 3(8)703-708 (1990) teach the introduction of a cysteine into an antibody (there the C<sub>H</sub>1 domain of a heavy chain) and the site-specific attachment of effector or reporter molecules through the introduced cysteine.

15 Despite the advances in *E. coli* expression of functional antibody fragments shown in the literature, there remains a need for efficient and economical techniques for the production of bivalent antibodies, particularly F(ab')<sub>2</sub> molecules, and for methods which permit the tailoring of bivalent and bispecific F(ab')<sub>2</sub> molecules. It would be desirable to produce stable Fab'-SH polypeptides which may be conveniently coupled *in vitro* to form bivalent Fv or F(ab')<sub>2</sub> molecules.

20 It is therefore an object of this invention to provide methods for the preparation of polypeptides comprising Fv domains, particularly Fab', Fab'-SH and F(ab')<sub>2</sub> antibody fragments, in or derived from bacterial cell culture in high yield.

It is a further object of this invention to provide methods for the efficient 25 preparation of homogenous bivalent and bispecific F(ab')<sub>2</sub> antibody fragments.

It is another object of this invention to provide Fab' antibody fragments having 30 at least one hinge region cysteine present as a free thiol (Fab'-SH). It is a related object to obviate the inherent problems in generating Fab'-SH from intact antibodies: differences in susceptibility to proteolysis and non-specific cleavage, low yield, as well as partial reduction which is not completely selective for the hinge disulfide bond(s). It is another object of the present invention to prevent intra-hinge disulfide bond formation without resorting to the use of highly toxic arsenite to chelate vicinal thiols, or other inefficient and undesirable methods.

Other objects, features, and characteristics of the present invention will become 35 apparent upon consideration of the following description and the appended claims.

Summary of the Invention

The principal embodiments of this invention are based on the surprising identification of cysteinyl free thiol in cysteinyl residues located outside of the light-heavy chain interface of recombinant microbial periplasmic antibody fragments, and the surprising discovery that Fv variants can be produced which contain only a single cysteinyl residue in the free thiol form. This facilitates the preparation of homogeneous recombinant F(ab')<sub>2</sub> and other Fv-containing bivalent polypeptides. Accordingly, in one embodiment this invention comprises expressing and secreting into the periplasm of a recombinant microbial cell culture a Fv polypeptide containing an immunoglobulin heavy chain Fv region and an immunoglobulin light chain Fv region, said light or heavy chain also comprising an unpaired cysteinyl residue as a free thiol, and recovering said polypeptide under conditions that substantially maintain said cysteinyl residue as the free thiol.

It will be understood that either of the Fv light or heavy chains optionally is fused to a polypeptide sequence which contains one or more cysteinyl residues, provided that at least one of such cysteinyl residues located in the domain C-terminal to either of the light or heavy chain Fv is present as the free thiol in the periplasm. Suitable polypeptide sequences include but are not limited to sequences derived from immunoglobulins, carrier proteins, receptors, growth factors, nutrient substances, cellular adhesive molecules, tissue-specific adhesion factors, enzymes, toxins and the like. Typically, an unpaired cysteinyl residue is substituted at the C-terminus of either the light or the heavy chain Fv domain or at the C-terminus of the light chain or heavy chain CH1 domain. The fused polypeptide sequence may comprise only the cysteinyl residue, or the cysteinyl residue can be present in a polypeptide fused to the C-terminus of (1) one of the Fv domains, (2) the light chain constant region or (3) the CH1 domain of a heavy chain. In the latter instance, the Fv-CH1 domain is fused to a hinge region bearing a single free thiol cysteinyl residue, and preferably a hinge region amino acid sequence variant containing only a single cysteinyl residue. However, any other sequence containing a single free thiol cysteinyl residue is useful in place of the immunoglobulin hinge region.

In other embodiments, the Fv-CH1 domain is fused to a hinge region bearing more than one, typically two or three, free thiol residues, preferably a hinge region amino acid sequence variant containing two or more cysteinyl residues. For example, Fab'-SH amino acid sequence variants which have been modified to contain three

cysteinyl residues may be produced recombinantly in a suitable host cell, for example CHO cells or *E. coli*, and are conveniently coupled *in vivo* by the host cell to form F(ab')<sub>2</sub> with three disulfide bonds connecting the heavy chains.

More particularly, the objects of this invention are accomplished by a method 5 for the production of a Fab' antibody polypeptide having at least one hinge region cysteine present as a free thiol (Fab'-SH), comprising the steps of:

- a. expressing nucleic acid encoding an immunoglobulin presequence comprising Fab' in a microbial host cell culture transformed with a vector comprising said nucleic acid operably linked to control sequences recognized by the host cell transformed 10 with the vector, under conditions suitable for the secretion of Fab' to the periplasmic space of the host cell and formation of Fab'-SH; and
- b. recovering Fab'-(SH)<sub>n</sub> from said host cell, where n is greater than or equal to one.

Additional embodiments of the method of this invention comprise the 15 subsequent recovery (including purification) of the Fab'-SH under conditions suitable for maintaining the hinge cysteinyl thiol(s) in protonated form. In certain embodiments, a metal ion chelating agent and/or a protease inhibitor is present during the culturing of the transformed cell or in the recovery of the Fab'-SH.

Optionally, the Fab'-SH is released from the host by freeze-thawing the host 20 cell, subjecting it to osmotic shock, preparing a cell paste and purifying the Fab'-SH from the cell paste. Optionally, release of Fab'-SH from the host cell is facilitated by enzymatic digestion of the cell e.g., using lysozyme or physical disruption, e.g., by sonication or by use of a French press.

The Fv or polypeptide fused to the Fv (typically the hinge sequence) is modified 25 in certain embodiments so that only one cysteinyl residue is present which, in the periplasm, exhibits free thiol. Thus, for example, the Fab amino acid sequence is modified by deleting or substituting all of the hinge sequence cysteinyl residues C-terminal to the first cysteine. Similarly, to obtain embodiments where, for example, three cysteinyl residues are to be present as the free thiol form, all of the hinge 30 sequence cysteinyl residues C-terminal to the third cysteine are modified by deletion or substitution.

This invention also provides Fv, Fab', Fab'-SH and F(ab')<sub>2</sub> polypeptides which do not have a disulfide bond between the light and heavy chains. These polypeptides are referred to herein as "linkless". For example, the amino acid sequence of the 35 polypeptide is modified by substituting or deleting one or, preferably, both of the two

cysteines which form the inter-chain (heavy-light) disulfide bond. Typically, these cysteines are replaced with serines, although it is within the scope hereof to covalently modify one or both of the cysteine side chain so as to be incapable of forming a disulfide bond. In certain embodiments, there is a very strong interaction 5 between the light and heavy chains such that eliminating the interchain disulfide does not result undesirable levels of dissociation between the light and heavy chains. The linkless Fab'-SH polypeptides advantageously allow homogeneous F(ab')<sub>2</sub> to be prepared by chemical coupling.

In another aspect, a method for the preparation of a polypeptide comprising 10 F(ab')<sub>2</sub> comprises the steps of:

a. expressing nucleic acid encoding an immunoglobulin presequence comprising a first Fab' in a microbial host cell culture transformed with a vector comprising the nucleic acid operably linked to control sequences recognized by the host cell transformed with the vector, under conditions suitable for the secretion of said first 15 Fab' to the periplasmic space of the host cell and formation of Fab'-SH, said first Fab' being capable of binding a first epitope;

b. expressing nucleic acid encoding an immunoglobulin presequence comprising a second Fab' in a microbial host cell culture transformed with a vector comprising the nucleic acid operably linked to control sequences recognized by the host cell transformed with the vector, under conditions suitable for the secretion of said second Fab' to the periplasmic space of the host cell and formation of Fab'-SH, said second Fab' being capable of binding a second epitope;

c. recovering said first and second Fab'-SH from said host cells; and  
d. forming covalent bonds between the free thiol cysteinyl residues of said first 25 and second Fab'-SH to form bivalent F(ab')<sub>2</sub>.

In embodiments of the invention, efficient directed disulfide bond formation occurs *in vitro*, utilizing directed covalent coupling methods known in the art whereby bispecific bivalent antibodies are produced. In particularly preferred embodiments, the disulfide bond formation between the first and second Fab'-SH comprises the 30 following steps:

a. reacting the first Fab'-SH with (i) 5,5'-dithiobis (2-nitrobenzoic acid) (DTNB) to form a thionitrobenzoate derivative Fab'-TNB or (ii) a bifunctional maleimide;  
b. directly coupling said first Fab'-TNB or maleimidated Fab' to the second Fab'-SH to form a F(ab')<sub>2</sub>; and  
35 c. recovering said F(ab')<sub>2</sub>.

Novel F(ab')<sub>2</sub> compositions are provided by the methods of this invention. Such compositions are

- a. essentially free of F(ab')<sub>2</sub> Fv regions containing cysteinyl residues with derivatized sulphhydryl groups, except for any native disulfide bond found in the native Fv regions,
- b. entirely free of F(ab')<sub>2</sub> having hinge region intrachain disulfide bonds,
- c. entirely free of contaminating arsenite, and
- d. entirely homogenous as to the heavy chain C-terminal amino acid residue.

In another embodiment, high level expression of immunoglobulins or fragments thereof in recombinant microbes is achieved by a method comprising culturing a host cell transformed with nucleic acid encoding an immunoglobulin polypeptide under the transcriptional control of an inducible promoter/operator system whereby expression of the polypeptide is repressed prior to induction sufficient to permit post-induction polypeptide levels in the cell culture of greater than about 1 gram of polypeptide per liter of cell culture. Typically, this is accomplished by using a powerful promoter such as phoA in a low copy number vector or in a host engineered to express levels of repressor sufficient to fully occupy all phoA operator sites and fully repress the operon.

20 Brief Description of the Drawings

FIGURE 1 shows the plasmid pA19 as used for the co-secretion of huMAb4D5-8  $\kappa$  light chain and heavy chain Fd' fragment from *E. coli*. The Fab' expression unit is dicistronic with both chains under the transcriptional control of the *PhoA* promoter (C. N. Chang *et al.*, *Gene* 44: 121 (1986)) which is inducible by phosphate starvation. The humanized variable domains (huV<sub>L</sub> and huV<sub>H</sub>, P. Carter *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 89(10):4285-9 (1992)) are precisely fused on their 5' ends to a gene segment encoding the heat stable enterotoxin II (stII) signal sequence (R. N. Picken *et al.*, *Infect. Immun.* 42: 269 (1983)) and on their 3' side to human  $\kappa_1$  (C<sub>L</sub>, W. Palm and N. Hilschmann, *Z. Physiol. Chem.* 356: 167 (1975)) and IgG1 (C<sub>H1</sub>, J. W. Ellison *et al.*, *Nucleic Acids Res.* 10: 4071 (1982)) constant domains respectively. The coding regions are separated by 83 base pairs and each is preceded by a ribosomal binding site (Picken *et al.*, *supra*) to enable efficient initiation of translation. The Fab' expression unit was cloned into the Eco RI site of pBR322 (F. Bolivar *et al.*, *Gene* 2: 95 (1977)) previously modified by removal of the Sal I and Sph I sites without changing the amino acid sequence of the tetracycline

-10-

resistance gene product. Construction of different Fab' variants was facilitated by installing unique Sal I and Sph I sites towards the end of the  $C_H1$  gene and immediately preceding the bacteriophage  $\lambda$   $t_0$  transcriptional terminator (S. Scholtissek, *et al.*, *Nucleic Acids Res.* 15: 3185 (1987)).

5

**FIGURE 2** shows the purification of the Fab, Fab' and  $F(ab')_2$  (Cys Ala Ala variant) fragments of huMAb4D5-8. Proteins were analyzed by SDS-PAGE on a 4 to 20% gel under non-reducing conditions with Coomassie brilliant blue (R250). Samples shown are protein molecular weight markers (lane 1), Fab (lane 2); Fab'-SH before (lane 3) and after (lane 4) a mock coupling reaction; Fab'-TNB before (lane 5) and after (lane 6) a mock coupling reaction; Fab'-SH coupled with Fab'-TNB (lane 7),  $F(ab')_2$  purified by gel filtration (lane 8) and  $F(ab')_2$  derived from limited pepsin digestion (E. Lamoyi, *et al.*, *Methods Enzymol.* 121: 652 (1986)) of full length huMAb4D5-8 expressed in mammalian cells (P. Carter *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* (1992) *supra* (lane 9)). Antibody fragments (2  $\mu$ g per sample) were reacted with 4 mM iodoacetamide prior to electrophoresis. The huMAb4D5-8 Fab fragment was purified from fermentation supernatants by centrifugation to remove cell debris, DEAE sepharose ion exchange chromatography and then affinity purified using protein A CL 4B or protein G sepharose. HuMAb4D5-8 Fab'-SH was purified by thawing 1.5 g cell paste in the presence of 21 ml of 100 mM sodium acetate (pH 3.5), 10 mM EDTA, 0.2 mM PMSF, 5  $\mu$ M pepstatin, 5  $\mu$ M leupeptin, 2.4 mM benzamidine. Cell debris was removed by centrifugation (40, 000g, 10 min, 4°C). The resultant supernatant (pH 5.0) was passed over DEAE sepharose and loaded on to a 2 ml protein G sepharose column. Protein eluted with 100 mM sodium acetate (pH 3.5), 10 mM EDTA was adjusted to pH 4.0 in the presence of 1.5M  $(NH_4)_2SO_4$  and loaded on to 2 ml phenyl Toyopearl column. Fab'-SH was eluted with 20 mM sodium acetate (pH 4.0), 0.8 M  $(NH_4)_2SO_4$ , 10 mM EDTA, reduced in volume by ultrafiltration (Centriprep-10, Amicon) and buffer exchanged into 10 mM sodium acetate (pH 5.0), 10 mM EDTA by G25 gel filtration. The Fab thionitrobenzoate derivative was prepared in a similar manner except that the DEAE flow through was adjusted to 5 mM DTNB and pH 7.5. The total concentration of huMAb4D5-8 Fab and Fab' variants was determined from the measured absorbance at 280 nm and the extinction coefficient determined by amino acid composition analysis ( $\epsilon^{0.1\%} = 1.56$ ). The free thiol content of huMAb4D5-8 Fab'-SH was determined by analysis with DTNB as described (T. E. Creighton, *Protein Structure, a Practical Approach* (IRL

Press, Oxford, UK, 1990), p. 157), whereas the TNB content of huMAb4D5-8 Fab'-TNB was determined by the yield upon reduction with dithiothreitol. Equimolar quantities of huMAb4D5-8 Fab'-TNB (by TNB content) and Fab'-SH (by -SH content) were coupled at a combined concentration of  $\geq 1.4$  mg/ml in the presence of 100 5 mM tris-HCl (pH 7.5) and 10 mM EDTA for 1 hour at 37°C. HuMAb4D5-8 F(ab')<sub>2</sub> was isolated from the coupling reaction by S100-HR gel filtration (Pharmacia) in the presence of phosphate-buffered saline. The F(ab')<sub>2</sub> samples were passed through a sterile 0.2  $\mu$ m filter and stored either at 4°C or flash frozen in liquid nitrogen and stored at -70°C.

10

#### Detailed Description of the Invention

##### Definitions

In general, the following words or phrases have the indicated definitions when 15 used in the description, examples, and claims:

The term Fv is defined to be a covalently or noncovalently-associated heavy and light chain heterodimer which does not contain constant domains.

The term Fab' is defined herein as a polypeptide comprising a heterodimer of the variable domain and the first constant domain of an antibody heavy chain, plus 20 the variable domain and constant domain of an antibody light chain, plus at least one additional amino acid residue at the carboxy terminus of the heavy chain C<sub>H</sub>1 domain including one or more cysteine residues. F(ab')<sub>2</sub> antibody fragments are pairs of Fab' antibody fragments which are linked by a covalent bond(s).

The Fab' heavy chain may include a hinge region. This may be any desired 25 hinge amino acid sequence. Alternatively the hinge may be entirely omitted in favor of one or more cysteine residues or, preferably a short (about 1-10 residues) cysteine-containing polypeptide. In certain applications, a common naturally occurring antibody hinge sequence (cysteine followed by two prolines and then another cysteine) is used; this sequence is found in the hinge of human IgG<sub>1</sub> molecules (E. A. 30 Kabat, *et al.*, *Sequences of Proteins of Immunological Interest* 3rd edition (National Institutes of Health, Bethesda, MD, 1987)). In other embodiments, the hinge region is selected from another desired antibody class or isotype. In certain preferred embodiments of this invention, the C-terminus of the C<sub>H</sub>1 of Fab' is fused to the sequence Cys X X. X preferably is Ala, although it may be any other residue such 35 as Arg, Asp, or Pro. One or both X amino acid residues may be deleted.

The "hinge region" is the amino acid sequence located between C<sub>H</sub>1 and C<sub>H</sub>2 in native immunoglobulins or any sequence variant thereof. In the case of the humanized 4D5 antibody described *infra*, the hinge region is located between residues 224 (asp in...Cys Asp Lys...) and 233 (Pro in ...Cys Pro Ala). Analogous regions of 5 other immunoglobulins will be employed, although it will be understood that the size and sequence of the hinge region may vary widely. For example, the hinge region of a human IgG1 is only about 10 residues, whereas that of human IgG<sub>3</sub> is about 60 residues.

The term Fv-SH or Fab'-SH is defined herein as a Fv or Fab' polypeptide having 10 at least one cysteinyl free thiol. Preferably the free thiol is in the hinge region, with the light and heavy chain cysteine residues that ordinarily participate in inter-chain bonding being present in their native form. In the most preferred embodiments of this invention, the Fab'-SH polypeptide composition is free of heterogenous 15 proteolytic degradation fragments. In certain embodiments, the Fab'-SH polypeptide is also substantially (greater than about 90 mole percent) free of Fab' fragments wherein heavy and light chains have been reduced or otherwise derivatized so as not to be present in their native state, e.g. by the formation of aberrant disulfides or sulfhydryl addition products. In alternative embodiments, the Fab'-SH has heavy and light chains which are not covalently coupled.

20 A humanized antibody for the purposes herein is an immunoglobulin amino acid sequence variant or fragment thereof which is capable of binding to a predetermined antigen and which comprises a FR region having substantially the amino acid sequence of a human immunoglobulin and a CDR having substantially the amino acid sequence of a non-human immunoglobulin or a sequence engineered to bind to a 25 preselected antigen.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, 30 optionally an operator sequence, a ribosome binding site, and transcriptional terminators. Particularly preferred are highly regulated inducible promoters that suppress Fab' polypeptide synthesis at levels below growth-inhibitory amounts while the cell culture is growing and maturing, for example, during the log phase.

Nucleic acid is "operably linked" when it is placed into a functional relationship 35 with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a

preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the

5 DNA sequences being linked are contiguous and, in the case of a secretory leader, contiguous and in same reading frame. However enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, then synthetic oligonucleotide adaptors or linkers are used in accord with conventional practice.

10 An "exogenous" element is defined herein to mean a nucleic acid sequence that is foreign to the cell, or homologous to the cell but in a position within the host cell nucleic acid in which the element is ordinarily not found.

15 As used herein, the expressions "cell" and "cell culture" are used interchangeably and all such designations include progeny. Thus, the words "transformants" and "transformed cells" include the primary subject cell and cultures derived therefrom without regard for the number of transfers. It is also understood that all progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. Mutant progeny that have the same function or biological activity as screened for in the originally transformed cell are included. Different 20 designations are will be clear from the contextually clear.

25 "Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are commercially available, are publicly available on an unrestricted basis, or can be constructed from such available plasmids in accord with published procedures. In addition, other equivalent plasmids are known in the art and will be apparent to the ordinary artisan.

30 "Recovery" or "isolation" of a given fragment of DNA from a restriction digest means separation of the digest on polyacrylamide or agarose gel by electrophoresis, identification of the fragment of interest by comparison of its mobility versus that of marker DNA fragments of known molecular weight, removal of the gel section containing the desired fragment, and separation of the gel from DNA. This procedure is known generally. For example, see *Lawn et al.*, Nucleic Acids Res., 9: 6103-6114 (1981), and *Goeddel et al.*, Nucleic Acids Res., 8: 4057 (1980).

35 "Preparation" of DNA from cells means isolating the plasmid DNA from a culture of the host cells. Commonly used methods for DNA preparation are the large and small scale plasmid preparations described in sections 1.25-1.33 of Sambrook

*et al.*, (Molecular Cloning: A Laboratory Manual New York: Cold Spring Harbor Laboratory Press, 1989). DNA preparations are purified by methods well known in the art (see section 1.40 of Sambrook *et al.*, *supra*).

5

Suitable Methods for Practicing the Invention

Typically a cultured microbial host cell is transformed with a vector comprising Fab'-encoding nucleic acid (i.e., nucleic acid encoding heavy chain Fd fragment and light chain) operably linked to control sequences recognized by the host cell 10 transformed with the vector (hereafter, "Fab" will be referred to as a specific embodiment, but it will be understood that any Fv-containing antibody fragments or fusion/sequence derivative thereof can be used provided that a cysteine residue capable of forming a periplasmic free thiol is present in the Fv region or sequence fused thereto). The cells are cultured under conditions suitable for the secretion of 15 Fab' into the periplasmic space of the host cell and formation of the free thiol. In general, a dicistronic operon is used to direct the co-expression of corresponding light and heavy chain fragments. Alternatively, separate chains are expressed from 20 separate promoters on the same or different plasmids. Secondly, the Fab chains are preceded by signal sequences to direct secretion into the periplasmic space, where it is believed that the redox environment favors disulfide bond formation for assembly 25 of the light and heavy chain fragments but not disulfide bond formation between hinge cysteine residues. In particularly preferred embodiments, the expression control sequence is the *E. coli* *phoA* promoter (C. N. Chang *et al.*, *Gene* 44: 121 (1986)) inducible by phosphate starvation and the signal sequence is the heat-stable enterotoxin II signal sequence (R. N. Picken *et al.*, *Infect. Immun.* 42: 269 (1983)).

It is currently preferred that the antibody fragments be expressed in bacterial cells grown at high cell density in a fermentor. Suitable fermentation conditions are described in the Example below.

The polypeptides containing free thiol are recovered from the fermentation 30 media and/or recovered from freeze-thawed cells (typically by osmotic shock) and subsequently purified. Recovery (including purification) is most successful if the Fab'-SH is maintained in the protonated form. Other conditions for maintaining the protonated form include the use of organic solvents or other agents for shifting the pKa of dissociation of -SH. This is conveniently accomplished at acid pH, i.e., 35 preferably 2 or more pH units below the pKa of the hinge or unpaired cysteinyl thiol.

-15-

Alternatively, the Fab'-SH is reacted with a protective group such as TNB or p-methoxybenzyl in order to maintain the Fab' in a homogenous state suitable for further reaction. In yet another alternative, pyridine disulfide is added to the Fab'-SH to form a mixed disulfide; this stabilizes the free sulphydryl until it is deprotected for 5 coupling or other processing. In still another alternative embodiment, the free sulphydryl is not protected but is reduced prior to coupling or further processing. Suitable protecting groups, known in the art, are described in E. Gross & J. Meiemhofer, *The Peptides: Analysis, Structure, Biology* Vol 3: Protection of Functional Groups in Peptide Synthesis (Academic Press, New York, 1981).

10 Affinity purification, such as on streptococcal protein G sepharose or staphylococcal protein A at acidic pH (typically, about pH 4 to pH 6, preferably about pH 5.0) is preferred. Alternatively, two-phase liquid extraction may be used. Small amounts of contaminating proteolytic fragments are readily removed by hydrophobic interaction chromatography, using, for example, silica gel and/or alkyl or aryl-substituted chromatography resins such as phenyl Toyopearl. It is preferred that a 15 cocktail of protease inhibitors be used (such as phenylmethylsulfonyl fluoride (PMSF), leupeptin, pepstatin and/or benzamidine) to inactivate or inhibit host proteases, and to include a metal ion chelating agent such as EDTA in the culture and recovery procedures. The chelating agent is selected and included in an amount to prevent 20 metal ion catalysis of disulfide bond formation.

In certain embodiments, Fab'-(SH)<sub>n</sub> (where n is equal to or greater than one) form F(ab')<sub>2</sub> *in vivo* during recombinant production. In these embodiments, the Fab'-(SH) amino acid sequence has been modified to contain preferably more than one cysteinyl residue. For example, F(ab')<sub>2</sub> having three disulfide bonds between heavy 25 chains are recovered directly from an *E. coli* cell paste, following the culture in *E. coli* of Fab'-SH<sub>3</sub>.

In other embodiments, F(ab')<sub>2</sub> is prepared from purified Fab'-(SH)<sub>n</sub> (where n is equal to or greater than one) by *in vitro* chemical coupling using cross-linking agents or adventitious oxidants such as dissolved oxygen. In the latter instance, purified 30 Fab' will form F(ab')<sub>2</sub> due simply to air oxidation. In addition, bispecific F(ab')<sub>2</sub> is produced by the use of Fab'-SH derivatives which direct the coupling of discrete Fab' and prevent the formation of bivalent monospecific F(ab')<sub>2</sub>. One suitable chemistry for mono- or bi-specific F(ab')<sub>2</sub> is that of Brennan *et al.*, *supra*. A Fab'-TNB derivative is prepared in a similar manner to Fab'-SH, except that Fab'-SH released from freeze-thawed cells by osmotic shock is adjusted to about pH 7.5 in the presence of excess 35

DTNB. Equimolar quantities of Fab'-SH and Fab'-TNB are coupled together efficiently to form the F(ab')<sub>2</sub> fragment by a disulfide exchange reaction in the presence of EDTA.

The linkless Fv, Fab', Fab'-SH and F(ab')<sub>2</sub> polypeptides of this invention do not have a disulfide bond between the light and heavy chains. Typically, the polypeptide is modified as the result of direct expression of a modified polypeptide, or by chemical or enzymatic means. Typical amino acid sequence modifications of the polypeptide involve substituting or deleting one or, preferably, both of the two cysteines which form the inter-chain (heavy-light) disulfide bond. Typically, these cysteines are replaced with serines, although it is within the scope hereof to covalently modify one or both of the cysteine side chain so as to be incapable of forming a disulfide bond.

One or both of the cysteines may be modified with fatty acids or other chemical groups and rendered incapable of forming a disulfide bond by a variety of known methods, including but not limited to the following methods: (1) covalent modification using dehydrating or activating agents such as N,N'-dicyclohexylcarbodiimide (DCC) or ethyoxy-ethoxycarbonyl-dihydroquinoline (EEDQ); (2) acylation, using ketenes, anhydrides, isothiocyanates, or beta-lactones; (3) carbamoylation using cyanates; (4) hemimercaptal or hemimercaptol formation using aldehydes and some keto acids; (5) alkylation and arylation, by addition to activated double bonds (using N-ethylmaleimide), by reactions with quinones, by reaction with haloacids and their amides (using iodoacetic acid, or alpha-bromo-hexadecanoic acid), by methylation reactions (such as with dimethylsulfate), by sulfoalkylation, by arylation (with nitrobenzene compounds), or by reaction with diazo compounds; (6) reaction with metal ions such as silver and organic mercury compounds such as mercuric chloride; (7) reaction with arsenic compounds; (8) reaction with sulfites; (9) oxidation reactions; and (10) reaction with sulfenyl halides to form mixed disulfides.

The polypeptide may be modified within a host cell, or as a post-translational modification to the recombinantly produced polypeptide. It is currently preferred that any post-translational modifications take place within 24, and preferably within a few hours of recovery from a host cell or microorganism.

Methods for detecting the presence of modifications to the cysteines of the polypeptides of this invention lung surfactant protein are commonly known, such as through the analysis of the mass spectra of a sample of the polypeptide. Alternatively, thin layer chromatography (TLC) is performed on a sample. While this

approach to peptide analysis is common in the field, see e.g. Stuart and Young, Solid Phase Peptide Synthesis, pp 103-107, and 118-122, (Pierce Chem. Co., 2d.ed., 1984).

5 In certain embodiments, there is a very strong interaction between the light and heavy chains such that eliminating the interchain disulfide does not result undesirable levels of dissociation between the light and heavy chains. The linkless Fab'-SH polypeptides advantageously allow homogeneous F(ab')<sub>2</sub> to be prepared by chemical coupling.

10 The Fab sequences of this invention are obtained from conventional sources of antibodies. Polyclonal antibodies to an antigen generally are raised in animals by multiple subcutaneous (sc) or intraperitoneal (ip) injections of the antigen and an adjuvant. It may be useful to conjugate the antigen or a fragment containing the target antigen amino acid sequence to a protein that is immunogenic in the species to be immunized, e.g., keyhole limpet hemocyanin, serum albumin, bovine 15 thyroglobulin, or soybean trypsin inhibitor using a bifunctional or derivatizing agent, for example, maleimidobenzoyl sulfosuccinimide ester (conjugation through cysteine residues), N-hydroxysuccinimide (through lysine residues), glutaraldehyde, succinic anhydride, SOCl<sub>2</sub>, or R<sup>1</sup>N = C = NR, where R and R<sup>1</sup> are different alkyl groups.

20 The route and schedule of immunization of the animal or cultured antibody-producing cells therefrom are generally in keeping with established and conventional techniques for antibody stimulation and production. While mice are frequently employed as the test model, it is contemplated that any mammalian subject or antibody-producing cells obtained therefrom can be employed.

25 After immunization, monoclonal antibodies are prepared by recovering immune lymphoid cells--typically spleen cells or lymphocytes from lymph node tissue--from immunized animals and immortalizing the cells in conventional fashion, e.g. by fusion with myeloma cells or by Epstein-Barr (EB)-virus transformation and screening for clones expressing the desired antibody. The hybridoma technique described originally by Köhler, G. and Milstein, C., *Nature* 256: 52-53 (1975) has been widely applied 30 to produce hybrid cell lines that secrete high levels of monoclonal antibodies against many specific antigens. Hybridomas secreting the desired antibody are identified by conventional methods. Routine methods are then employed to obtain DNA from hybridomas that encode the heavy and light chains of the selected antibody. Alternatively, one extracts antibody-specific messenger RNAs from B-cells of an 35 immunized animal, reverse transcribes these into complementary DNA (cDNA), and

amplifies the cDNA by PCR or by cloning it in a bacterial expression system. Another technique suitable for obtaining source heavy and light chain sequences uses a bacteriophage lambda vector system (which contains a leader sequence that secretes the expressed Fab protein into the periplasmic space) together with the generation 5 and screening of great numbers of functional antibody fragments for those which demonstrate the desired activity. This system is commercially available.

#### Amino Acid Sequence Variants

This invention also encompasses amino acid sequence variants of the native 10 Fab polypeptide sequences. These variants are prepared by introducing appropriate nucleotide changes into the DNA encoding the Fab or by *in vitro* synthesis of the desired Fab. Such variants include, for example, humanized variants of non-human 15 antibodies, as well as deletions from, or insertions or substitutions of, residues within particular amino acid sequences. Any combination of deletion, insertion, and 20 substitution can be made to arrive at the final construct, provided that the final construct possesses the desired characteristics. The amino acid changes also may alter post-translational processing of the target polypeptide, such as changing the number or position of glycosylation sites, introducing a membrane anchoring sequence into the constant domain or modifying the leader sequence of the native Fab.

In designing amino acid sequence variants of target polypeptides, the location 25 of the mutation site and the nature of the mutation will depend on the target polypeptide characteristic(s) to be modified. The sites for mutation can be modified individually or in series, e.g., by (1) substituting first with conservative amino acid choices and then with more radical selections depending upon the results achieved, (2) deleting the target residue, or (3) inserting residues of the same or a different 30 class adjacent to the located site, or combinations of options 1-3.

A useful method for identification of certain residues or regions of the target 35 polypeptide that are preferred locations for mutagenesis is called "alanine scanning mutagenesis" as described by Cunningham and Wells (*Science*, **244**: 1081-1085 [1989]), and Duncan, A.R. and Winter, G. (*Nature*, **322**: 738-740 [1988]). Here, a residue or group of target residues are identified (e.g., charged residues such as arg, asp, his, lys, and glu) and replaced by a neutral or negatively charged amino acid (most preferably alanine or polyalanine) to affect the interaction of the amino acids with the surrounding aqueous environment in or outside the cell. Thus, while the site

for introducing an amino acid sequence variation is predetermined, the nature of the mutation *per se* need not be predetermined. For example, to optimize the performance of a mutation at a given site, Ala scanning or random mutagenesis may be conducted at the target codon or region and the expressed target polypeptide variants are screened for the optimal combination of desired activity.

There are two principal variables in the construction of amino acid sequence variants: the location of the mutation site and the nature of the mutation. In general, the location and nature of the mutation chosen will depend upon the target polypeptide characteristic to be modified.

10 Amino acid sequence insertions include amino- and/or carboxyl-terminal fusions ranging in length from one residue to polypeptides containing a hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions (i.e., insertions within the target polypeptide sequence) may range generally from about 1 to 10 residues, more preferably 1 to 5, most preferably 1 to 3. Examples of terminal insertions include fusion of a heterologous N-terminal signal sequence to the N-terminus of the Fv polypeptide to facilitate the secretion of the mature Fv polypeptide from recombinant host cells. Such signal sequences generally will be obtained from, and thus homologous to, the intended host cell species. Suitable sequences for *E. coli* include STII or Ipp.

20 Other insertional variants of the target polypeptide include the fusion to the N- or C-terminus of the target polypeptide of immunogenic polypeptides, e.g., bacterial polypeptides such as beta-lactamase or an enzyme encoded by the *E. coli trp* locus, or yeast protein, and C-terminal fusions with proteins having a long half-life such as immunoglobulin constant regions (or other immunoglobulin regions), albumin, or ferritin, as described in WO 89/02922 published 6 April 1989. Additional suitable polypeptide sequences include but are not limited to sequences derived from immunoglobulins, carrier proteins, receptors, growth factors, nutrient substances, cellular adhesive molecules, tissue-specific adhesion factors, enzymes, toxins and the like. These examples will be in addition to insertions of cysteine or cysteine-containing polypeptides such as hinge regions which provide free thiol cysteinyl.

30 Another group of variants are amino acid substitution variants. These variants have at least one amino acid residue in the immunoglobulin polypeptide removed and a different residue inserted in its place. The sites of greatest interest for substitutional mutagenesis include the CDRs, FR and hinge regions. They include substitutions of cysteine for other residue and insertions which are substantially

different in terms of side-chain bulk, charge, and/or hydrophobicity. Other sites for substitution are described infra, considering the effect of the substitution of the antigen binding, affinity and other characteristics of a particular target antibody.

Fab itself is a deletional variant of intact immunoglobulin in which the heavy chain constant domain downstream from C<sub>H</sub>1 is deleted. Further, in preferred embodiments the C<sub>H</sub>1 domain is followed C-terminally by a cysteine-containing sequence such as Cys Ala Ala, or sequences having more than one Cys.

DNA encoding Fab amino acid sequence variants is prepared by a variety of methods known in the art. These methods include, but are not limited to, isolation from a natural source (in the case of naturally occurring amino acid sequence variants) or preparation by oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis, and cassette mutagenesis of an earlier prepared variant or a non-variant version of the target polypeptide or by total gene synthesis. These techniques may utilize target polypeptide nucleic acid (DNA or RNA), or nucleic acid complementary to the target polypeptide nucleic acid. Oligonucleotide-mediated mutagenesis is a preferred method for preparing substitution, deletion, and insertion variants of target polypeptide DNA.

PCR mutagenesis is also suitable for making amino acid variants of the Fab polypeptide. While the following discussion refers to DNA, it is understood that the technique also finds application with RNA. The PCR technique generally refers to the following procedure (see Erlich, *supra*, the chapter by R. Higuchi, p. 61-70): When small amounts of template DNA are used as starting material in PCR, primers that differ slightly in sequence from the corresponding region in a template DNA can be used to generate relatively large quantities of a specific DNA fragment that differs from the template sequence only at the positions where the primers differ from the template. For introduction of a mutation into a plasmid DNA, one of the primers is designed to overlap the position of the mutation and to contain the mutation; the sequence of the other primer must be identical to a stretch of sequence of the opposite strand of the plasmid, but this sequence can be located anywhere along the plasmid DNA. It is preferred, however, that the sequence of the second primer is located within 200 nucleotides from that of the first, such that in the end the entire amplified region of DNA bounded by the primers can be easily sequenced. PCR amplification using a primer pair like the one just described results in a population of DNA fragments that differ at the position of the mutation specified by the primer, and possibly at other positions, as template copying is somewhat error-prone.

If the ratio of template to product material is extremely low, the vast majority of product DNA fragments incorporate the desired mutation(s). This product material is used to replace the corresponding region in the plasmid that served as PCR template using standard DNA technology. Mutations at separate positions can be 5 introduced simultaneously by either using a mutant second primer, or performing a second PCR with different mutant primers and ligating the two resulting PCR fragments simultaneously to the vector fragment in a three (or more)-part ligation.

Another method for preparing variants, cassette mutagenesis, is based on the technique described by Wells *et al.* (*Gene*, 34: 315 [1985]). The starting material is 10 the plasmid (or other vector) comprising the Fab DNA to be mutated. The codon(s) in the Fab polypeptide DNA to be mutated are identified. There must be a unique restriction endonuclease site on each side of the identified mutation site(s). If no such restriction sites exist, they may be generated using the above-described oligonucleotide-mediated mutagenesis method to introduce them at appropriate 15 locations in the Fab polypeptide DNA. After the restriction sites have been introduced into the plasmid, the plasmid is cut at these sites to linearize it. A double-stranded oligonucleotide encoding the sequence of the DNA between the restriction sites but containing the desired mutation(s) is synthesized using standard procedures. The two strands are synthesized separately and then hybridized together 20 using standard techniques. This double-stranded oligonucleotide is referred to as the cassette. This cassette is designed to have 3' and 5' ends that are compatible with the ends of the linearized plasmid, such that it can be directly ligated to the plasmid. This plasmid now contains the mutated Fab polypeptide DNA sequence.

25 Insertion of DNA into a Vector

The cDNA or genomic DNA encoding the Fab polypeptide is inserted into a replicable vector for further cloning (amplification of the DNA) or for expression. Many vectors are available, and selection of the appropriate vector will depend on 1) 30 whether it is to be used for DNA amplification or for expression of the encoded protein, 2) the size of the DNA to be inserted into the vector, and 3) the host cell to be transformed with the vector. Each vector contains various components depending on its function (amplification of DNA or expression of DNA) and the host cell for which it is compatible. The vector components generally include, but are not limited to, one or more of the following: a signal sequence, an origin of replication, one or 35 more marker genes, a promoter, and a transcription termination sequence.

(a) Signal Sequence Component

In general, the signal sequence may be a component of the vector, or it may be a part of the target polypeptide DNA that is inserted into the vector. Included within the scope of this invention are Fab polypeptides with any native signal sequence deleted and replaced with a heterologous signal sequence. The heterologous signal sequence selected should be one that is recognized and processed (i.e. cleaved by a signal peptidase) by the host cell. For prokaryotic host cells that do not recognize and process the native Fab polypeptide signal sequence, the signal sequence is substituted by a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders.

(b) Origin of Replication Component

Expression and cloning vectors may, but need not, contain a nucleic acid sequence that enables the Fab nucleic acid to replicate in one or more selected host cells. Generally, in cloning vectors this sequence is one that enables the vector to replicate independently of the host chromosomal DNA, and includes origins of replication or autonomously replicating sequences. Such sequences are well known for a variety of microbes. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria.

DNA may also be replicated by insertion into the host genome. This is readily accomplished using *Bacillus* species as hosts, for example, by including in the vector a DNA sequence that is complementary to a sequence found in *Bacillus* genomic DNA. Transfection of *Bacillus* with this vector results in homologous recombination with the genome and insertion of the target polypeptide DNA. However, the recovery of genomic DNA encoding the target polypeptide is more complex than that of an exogenously replicated vector because restriction enzyme digestion is required to excise the target polypeptide DNA. Similarly, DNA also can be inserted into the genome of vertebrate and mammalian cells by conventional methods.

(c) Selection Gene Component

Expression and cloning vectors should contain a selection gene, also termed a

-23-

selectable marker. This gene encodes a protein necessary for the survival or growth of transformed host cells grown in a selective culture medium. Host cells not transformed with the vector containing the selection gene will not survive in the culture medium. Typical selection genes encode proteins that (a) confer resistance 5 to antibiotics or other toxins, e.g. ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g. the gene encoding D-alanine racemase for *Bacilli*.

One example of a selection scheme utilizes a drug to arrest growth of a host 10 cell. Those cells that are successfully transformed with a heterologous gene express a protein conferring drug resistance and thus survive the selection regimen. Examples of such dominant selection use the drugs neomycin (Southern *et al.*, J. Molec. Appl. Genet., 1: 327 [1982]), mycophenolic acid (Mulligan *et al.*, Science, 209: 1422 [1980]) or hygromycin (Sugden *et al.*, Mol. Cell. Biol., 5: 410-413 15 [1985]). The three examples given above employ bacterial genes under eukaryotic control to convey resistance to the appropriate drug G418 or neomycin (geneticin), xgpt (mycophenolic acid), or hygromycin, respectively.

(d) Promoter Component

20 Expression and cloning vectors will usually contain a promoter that is recognized by the host organism and is operably linked to the Fab polypeptide nucleic acid. Promoters are untranslated sequences located upstream (5') to the start codon of a the Fab structural gene (generally within about 100 to 1000 bp) that control its transcription and translation. Such promoters typically fall into two classes, inducible 25 and constitutive. Inducible promoters are promoters that initiate increased levels of transcription from DNA under their control in response to some change in culture conditions, e.g. the presence or absence of a nutrient or a change in temperature.

Inducible promoters under high regulation are preferred for the microbial expression of Fv-containing polypeptides. At this time a large number of promoters 30 recognized by a variety of potential host cells are well known. These promoters are operably linked to DNA encoding the Fab polypeptide by removing the promoter from the source DNA by restriction enzyme digestion and inserting the isolated promoter sequence into the vector. Both the native Fab polypeptide promoter sequence and many heterologous promoters may be used to direct amplification and/or expression 35 of the Fab polypeptide DNA. However, heterologous promoters are preferred, as they

-24-

generally permit greater transcription and higher yields of expressed target polypeptide as compared to the native target polypeptide promoter.

Promoters suitable for use with prokaryotic hosts include the  $\beta$ -lactamase and lactose promoter systems (Chang *et al.*, *Nature*, 275: 615 [1978]; and Goeddel *et al.*, *Nature*, 281: 544 [1979]), alkaline phosphatase, a tryptophan (trp) promoter system (Goeddel, *Nucleic Acids Res.*, 8: 4057 [1980] and EP 36,776) and hybrid promoters such as the tac promoter (deBoer *et al.*, *Proc. Natl. Acad. Sci. USA*, 80: 21-25 [1983]). However, other known bacterial promoters are suitable. Their nucleotide sequences have been published, thereby enabling a skilled worker operably to ligate them to DNA encoding the target polypeptide (Siebenlist *et al.*, *Cell*, 20: 269 [1980]) using linkers or adaptors to supply any required restriction sites. Promoters for use in bacterial systems also generally will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding the target polypeptide.

Construction of suitable vectors containing one or more of the above listed components employs standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to generate the plasmids required.

#### Selection and Transformation of Host Cells

Suitable host cells for expressing Fab are microbial cells such as yeast, fungi, and prokaryotes. Suitable prokaryotes include eubacteria, such as Gram-negative or Gram-positive organisms, for example, *E. coli*, *Bacilli* such as *B. subtilis*, *Pseudomonas* species such as *P. aeruginosa*, *Salmonella typhimurium*, or *Serratia marcescans*. One preferred *E. coli* cloning host is *E. coli* 294 (ATCC 31,446), although other strains such as *E. coli* B, *E. coli*  $\lambda$ 1776 (ATCC 31,537), *E. coli* RV308(ATCC 31,608) and *E. coli* W3110 (ATCC 27,325) are suitable. These examples are illustrative rather than limiting. Preferably the host cell should secrete minimal amounts of proteolytic enzymes, and additional protease inhibitors may desirably be incorporated in the cell culture.

Host cells are transfected and preferably transformed with the above-described expression or cloning vectors of this invention and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

Culturing the Host Cells

Cells used to produce the Fab polypeptide of this invention are cultured in suitable media as described generally in Sambrook *et al.*, (Molecular Cloning: A Laboratory Manual New York: Cold Spring Harbor Laboratory Press, 1989). Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art. The culture conditions, such as temperature, pH, and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

It is currently preferred that the bacterial host cells be cultured at temperatures from 37 °C to 29 °C, although temperatures as low as 20 °C may be suitable. Optimal temperatures will depend on the host cells, the Fab sequence and other parameters. 37°C is generally preferred.

15 Purification of Fab polypeptides

Soluble polypeptides are recovered from recombinant cell culture to obtain preparations that are substantially homogeneous as to Fab. As a first step, the culture medium or periplasmic preparation is centrifuged to remove particulate cell debris. Periplasmic preparations are obtained in conventional fashion, e.g. by freeze-thaw or osmotic shock methods. The membrane and soluble protein fractions are then separated. The Fab polypeptide is then purified from the soluble protein fraction. The following procedures are exemplary of suitable purification procedures: fractionation on immunoaffinity or ion-exchange columns; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A or protein G affinity matrix (e.g. Sepharose) columns; and hydrophobic interaction chromatography.

Fab polypeptide variants in which residues have been deleted, inserted or substituted are recovered in the same fashion, taking account of any substantial changes in properties occasioned by the variation. For example, preparation of a Fab polypeptide fusion with another protein or polypeptide, e.g. a bacterial or viral antigen, facilitates purification since an immunoaffinity column containing antibody to the antigen is used to adsorb the fusion. Immunoaffinity columns such as a rabbit polyclonal anti-target polypeptide column can be employed to absorb the target

polypeptide variant by binding it to at least one remaining immune epitope. A protease inhibitor also is useful to inhibit proteolytic degradation during purification, and antibiotics may be included to prevent the growth of adventitious contaminants.

5

#### Utility of the Fab-containing Polypeptides

The antibody fragments of this invention are useful in diagnostic assays for antigen in specific cells, fluids or tissues, for immunoaffinity purification of the antigens and for therapies which are based on antigen antagonism.

Analytical methods for the antigen bound by the Fab polypeptide are conventional and may use a label bound to the Fab. The label used with the Fab polypeptide is any detectable functionality that does not interfere with its binding to Fab. Numerous labels are known, including the radioisotopes  $^{32}\text{P}$ ,  $^{32}\text{S}$ ,  $^{14}\text{C}$ ,  $^{125}\text{I}$ ,  $^{3}\text{H}$ , and  $^{131}\text{I}$ , fluorophores such as rare earth chelates or fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, luciferases, e.g., firefly luciferase and bacterial luciferase (U.S. Pat. No. 4,737,456), luciferin, 2,3-dihydropthalazinediones, horseradish peroxidase (HRP), alkaline phosphatase,  $\beta$ -galactosidase, glucoamylase, lysozyme, saccharide oxidases, e.g., glucose oxidase, galactose oxidase, and glucose-6-phosphate dehydrogenase, heterocyclic oxidases such as uricase and xanthine oxidase, lactoperoxidase, biotin/avidin, spin labels, bacteriophage labels, stable free radicals, imaging radionuclides (such as Technetium) and the like.

Conventional methods are available to bind these labels covalently to proteins or polypeptides. For instance, coupling agents such as dialdehydes, carbodiimides, dimaleimides, bis-imides, bis-diazotized benzidine, and the like may be used to tag the antibodies with the above-described fluorescent, chemiluminescent, and enzyme labels. See, for example, U.S. Pat. Nos. 3,940,475 (fluorimetry) and 3,645,090 (enzymes); Hunter *et al.*, *Nature*, 144: 945 (1962); David *et al.*, *Biochemistry*, 13: 1014-1021 (1974); Pain *et al.*, *J. Immunol. Methods*, 40: 219-230 (1981); and Nygren, *J. Histochem. and Cytochem.*, 30: 407-412 (1982). Preferred labels herein are enzymes such as horseradish peroxidase and alkaline phosphatase. The conjugation of such label, including the enzymes, to the Fab-containing polypeptide is a standard manipulative procedure for one of ordinary skill in immunoassay techniques. See, for example, O'Sullivan *et al.*, "Methods for the Preparation of

Enzyme-antibody Conjugates for Use in Enzyme Immunoassay," in Methods in Enzymology, ed. J.J. Langone and H. Van Vunakis, Vol. 73 (Academic Press, New York, New York, 1981), pp. 147-166. Such bonding methods are suitable for use with the Fab polypeptides of this invention.

5        The Fab-containing polypeptides also may comprise an immunotoxin. For example, the Fab heavy chain is optionally conjugated to a cytotoxin such as ricin for use in AIDS therapy. Alternatively, the toxin may be a cytotoxic drug or an enzymatically active toxin of bacterial, fungal, plant or animal origin, or an enzymatically active fragment of such a toxin. Enzymatically active toxins and  
10      fragments thereof include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin,  
15      restrictocin, phenomycin, enomycin and the trichothecenes. In another embodiment, the antibodies are conjugated to nonpeptidyl drugs such as cis-platin or 5FU. Conjugates of the monoclonal antibody and such cytotoxic moieties are made using a variety of bifunctional protein coupling agents. Examples of such reagents are SPDP, IT, bifunctional derivatives of imidoesters such as dimethyl adipimidate HCl,  
20      active esters such as disuccinimidyl suberate, aldehydes such as glutaraldehyde, bis-azido compounds such as bis (p-azidobenzoyl) hexanediamine, bis-diazonium derivatives such as bis- (p-diazoniumbenzoyl)-ethylenediamine, diisocyanates such as tolylene 2,6-diisocyanate and bis-active fluorine compounds such as 1,5-difluoro-2,4-dinitrobenzene. The lysing portion of a toxin may be joined to the Fab antibody  
25      fragment.

Immunotoxins can be made in a variety of ways, as discussed herein. Commonly known crosslinking reagents can be used to yield stable conjugates.

When used *in vivo* for therapy, the Fab-containing fragments of the subject invention are administered to the patient in therapeutically effective amounts (i.e. 30 amounts that have desired therapeutic effect) in the same fashion as intact immunoglobulins. The products prepared in accord with the methods of this invention offer the advantage of substantial molecular homogeneity and are devoid of toxic contaminants heretofore used in preparing  $F(ab')_2$ .

The antibody compositions used in therapy are formulated and dosages 35 established in a fashion consistent with good medical practice taking into account the

-28-

disorder to be treated, the condition of the individual patient, the site of delivery of the composition, the method of administration and other factors known to practitioners. The antibody compositions are prepared for administration according to the description of preparation of polypeptides for administration, *infra*.

5

Example: Expression of active Fab, Fab', and F(ab'), antibody fragments

Overexpression of the *HER2* proto-oncogene product (p185<sup>HER2</sup>) has been associated with a variety of aggressive human malignancies. An *Escherichia coli* expression system has been developed that secretes functional Fab and Fab' fragments of a humanized antibody, huMAb4D5-8, at titers of about 1 to in excess of about 2 grams per liter as judged by binding to antigen, p185<sup>HER2</sup>. The Fab' fragment was recovered with the single hinge region cysteine present mainly as the free thiol (up to about 90 mole %) permitting efficient directed disulfide bond formation *in vitro* to form the bivalent F(ab')<sub>2</sub> antibody fragment. This molecule is indistinguishable from F(ab')<sub>2</sub> derived from proteolysis of intact antibody in antigen binding affinity and in anti-proliferative activity against the human breast tumor cell line, SK-BR-3, which over-expresses p185<sup>HER2</sup>, but unlike the proteolytic product, the F(ab')<sub>2</sub> here is C-terminally homogenous. This invention facilitates the construction of monospecific and bispecific F(ab')<sub>2</sub> antibody fragments, including naturally derived or humanized antibody fragments for research and therapeutic purposes. This invention is particularly applicable for developing antibody fragments capable of directing diagnostic or therapeutic moieties to target antigens such as tumor foci characterized by overexpression of the *HER2* proto-oncogene.

25

The murine monoclonal antibody known as muMAb4D5 (Fendly, B. M. *et al.*, *Cancer Res.* 50:1550-1558 (1990)) is directed against the extracellular domain (ECD) of p185<sup>HER2</sup>. The muMAb4D5 and its uses are described in copending PCT application WO 89/06692 published 27 July 1989. This murine antibody was deposited with the ATCC and designated ATCC CRL 10463. In this description, the terms muMAb4D5, chMAb4D5 and huMAb4D5 represent murine, chimerized and humanized versions of the monoclonal antibody 4D5, respectively.

30

muMAb4D5 has potential for clinical intervention in that it specifically inhibits the growth of human tumor cell lines overexpressing p185<sup>HER2</sup> (2, 3) and is rapidly internalized by target cells (data not shown). Amplification and / or overexpression of *HER2* is associated with multiple human malignancies and appears to be integrally

35

-29-

involved in progression of 25-30% of primary human breast and ovarian cancers (5). The muMAb4D5 molecule was previously "humanized" (6) in an attempt to improve its clinical efficacy by reducing immunogenicity and allowing it to support human effector functions. The humanized antibody, huMAb4D5-8, contains essentially only 5 the antigen binding loops from the murine parent antibody together with human variable region framework residues and constant domains. This humanized antibody has 3-fold higher affinity for p185<sup>HER2</sup> ECD than the murine parent, comparable anti-proliferative activity against p185<sup>HER2</sup> overexpressing tumor cells and supports antibody-dependent cellular cytotoxicity.

10 It was desired to enlarge the repertoire of anti-p185<sup>HER2</sup> antibody reagents to include the Fab and F(ab')<sub>2</sub> antibody fragments of huMAb4D5-8. The smaller size of these fragments compared to the intact antibody is likely to enhance their specific localization to solid tumors by improving tumor penetration and promoting more rapid removal from serum (reviewed in ref. 8). huMAb4D5-8 Fab and F(ab')<sub>2</sub> fragments 15 were obtained by direct recombinant expression to facilitate engineering of these fragments. Furthermore this strategy provided more homogeneous preparations of antibody fragments than could be obtained by limited proteolysis and partial reduction of intact antibodies (reviewed in ref. 9). Functional Fv and Fab fragments had been secreted from *E. coli* (10).

20 The strategy here for the *E. coli* secretion of antibody fragments (Fig. 1) shares two basic similarities with the work of others (10). Firstly a dicistronic operon is used to direct the co-expression of corresponding light and heavy chain fragments. Secondly the antibody chains are preceded by bacterial signal sequences to direct secretion into the periplasmic space of *E. coli* where the redox environment favors 25 disulfide bond formation and the light and heavy chain fragments may assemble. The system here differs from earlier strategies in three basic ways. Firstly the transcription unit utilizes the a highly regulated promoter, the *E. coli* *PhoA* promoter (11) inducible by phosphate starvation, and heat-stable enterotoxin II signal sequence (12). Secondly, the gene segment for the light chain precedes that for the heavy 30 chain Fd fragment (V<sub>H</sub> and C<sub>H</sub>1 domains). Thirdly, in order to express the Fab' fragment of huMAb4D5-8 the C<sub>H</sub>1 gene segment was extended to encode part of the cysteine-containing antibody hinge region. The sequence Cysteine followed by two Prolines and another Cysteine (CPC terminus) was initially chosen since it is found in the hinge region of human IgG<sub>1</sub> molecules (17) including the full length version of 35 huMAb4D5-8 (6). The construction of additional Fab' variants by cassette

-30-

mutagenesis (18) of the pBR322-based expression vector was facilitated by installing unique Sal I and Sph I restriction sites towards the end of the  $C_H1$  gene segment and immediately 3' to the stop codon, respectively.

The huMAb4D5-8 Fab fragment was expressed in a phage resistant derivative of *E. coli* RV308 (19) grown at high cell density in a fermentor (20). The titer of functional huMAb4D5-8 Fab in fermentation media is routinely 1 to 2 grams per liter as judged by  $p185^{HER2}$  ECD binding ELISA (6). Modest amounts of huMAb4D5-8 Fab (usually < 200 mg/l) are found associated with the cell paste and may be released by osmotic shock. The consensus framework region of the huMAb4D5-8 Fab fragment was found to bind tightly to both staphylococcal protein A and to streptococcal protein G, allowing their use for affinity purification. Very similar estimates of titers are obtained for culture media or cell paste samples after affinity purification on protein A prior to antigen-binding ELISA. Similar expression titers of  $p185^{HER2}$  ECD binding activity have been obtained for the huMAb4D5-8 Fab variant (cysteine, two prolines and another cysteine) and the additional Fab' variant described below.

Formation of  $F(ab')_2$  molecules requires chance encounters of Fab' hinge cysteinyl thiols to form disulfide bonds without the assistance of extensive interactions between  $C_H3$  domains possible in the case of intact antibodies. Thus high level expression of Fab' in the periplasmic space of *E. coli* was anticipated to drive formation of  $F(ab')_2$  *in vivo*. In fact < 10% of the Fab' molecule having the CPC terminus (isolated from either media or cell paste) were recovered as the bivalent form as judged by SDS-PAGE analysis after protein A purification. High resolution mass spectrometry and other studies suggested that substantial formation of an intramolecular disulfide bond between the two hinge cysteine residues had occurred. This possibility was precluded by the construction of an additional Fab' variant with a single hinge cysteine residue having the C-terminal sequence, Cys Ala Ala. Negligible quantities of  $F(ab')_2$  are formed when this Fab' variant is secreted from *E. coli* and DTNB analysis does not detect free thiol.

The Fab' molecules are recovered under conditions that maintain the hinge cysteine present as the free thiol and then readily and efficiently formed into  $F(ab')_2$  by directed coupling *in vitro*. For example, Brennan *et al.* (23) reacted a Fab' free thiol (Fab'-SH) with DTNB to form the thionitrobenzoate derivative (Fab'-TNB) which was then coupled to a second Fab'(Fab'-SH) to form the bispecific  $F(ab')_2$ . The following strategy allows the routine purification of intact functional huMAb4D5-8

-31-

Fab' Cys Ala Ala variant with 75-90% of the molecules containing a free hinge thiol as judged by DTNB analysis (Fig. 2): firstly, growth conditions were judiciously modified (20) to target secretion of Fab' to the periplasmic space of *E. coli* rather than into the culture media where the Fab' hinge thiol was found to be quantitatively 5 and covalently blocked. Secondly, the Fab' fragment was isolated from the cell paste and affinity purified on protein G sepharose at low pH (pH 5.0) to maintain the cysteinyl thiol in the less reactive protonated form. Thirdly, EDTA was added to chelate metal ions capable of catalyzing disulfide bond formation and to inactivate metallo proteases. Finally a cocktail of additional protease inhibitors 10 (phenylmethylsulfonyl fluoride (PMSF), leupeptin, pepstatin and benzamidine) virtually eliminated proteolysis of Fab' by *E. coli* proteases during the purification. Small amounts of contaminating proteolytic fragments in Fab' preparations were readily removed by hydrophobic interaction chromatography. The Fab'-TNB derivative was prepared in a similar manner except that Fab'-SH released from freeze-thawed cells 15 by osmotic shock was adjusted to neutral pH in the presence of excess DTNB.

Equimolar quantities of Fab'-SH and Fab-TNB were coupled together to efficiently form the  $F(ab')_2$  antibody fragment by a disulfide exchange reaction (Fig. 2). The coupling reaction was followed by monitoring the increase in absorbance at 20 412 nm upon release of the thionitrobenzoate anion and found to reach completion after 30 to 60 min at 37°C. No free thiol was detected in the reaction mix after coupling and the amount of remaining Fab' is consistent with that amount of unreactive material in the Fab'-SH and Fab'-TNB preparations.  $F(ab')_2$  was separated from Fab' by gel filtration on a S100-HR sizing column. Only trace quantities of huMAb4D5-8  $F(ab')_2$  were formed in mock coupling reactions containing either 25 Fab'-SH or Fab'-TNB alone, as anticipated (23).

The physical and chemical integrity of purified huMAb4D5-8 Fab and  $F(ab')_2$  was evaluated by SDS-PAGE (Fig. 2), analysis of the amino terminal sequence, amino acid composition plus free thiol content, and by circular dichroism. Purified huMAb4D5-8 Fab ( $M_r = 47.7$  kdal) and  $F(ab')_2$  ( $M_r = 96.0$  kdal) fragments analyzed 30 by SDS-PAGE under non-reducing conditions each gave a single major band of the expected mobility. After SDS-PAGE under reducing conditions both Fab and  $F(ab')_2$  antibody fragments gave a doublet of bands of similar intensities as expected from release of stoichiometric quantity of free light chain (23.4 kdal) and heavy chain Fd 35 (24.3 kdal) or Fd' (24.6 kdal) fragments (not shown). Amino terminal sequence analysis (8 cycles) of Fab and  $F(ab')_2$  antibody fragments gave the expected mixed

-32-

sequence from a stoichiometric 1:1 mixture of light and heavy chains ( $V_L$  /  $V_H$ : Asp/Glu, Ile/Val, Gln/Gln, Met/Leu, Thr/Val, Gln/Glu, Ser/Ser, and Pro/Gly with no evidence of additional sequences. No free thiol was detected in either Fab or  $F(ab')_2$  preparations by DTNB analysis as expected. Amino acid analysis (27) of acid hydrolysed Fab or  $F(ab')_2$  was in excellent agreement with the expected composition (6). The circular dichroism spectrum of the Fab fragment is characteristic of an immunoglobulin fold.

The function of huMAb4D5-8 Fab and  $F(ab')_2$  antibody fragments was investigated by measuring the binding affinity for the p185<sup>HER2</sup> ECD and by investigating their effect upon the proliferation of the p185<sup>HER2</sup> overexpressing human breast carcinoma line, SK-BR-3 (Table 1).

-33-

Table 1. Analysis of huMAb4D5-8 fragments by p185<sup>HER2</sup> ECD binding affinity and anti-proliferative activity with breast carcinoma, SK-BR-3 cells.

5	<u>huMAb4D5-8 variant</u>	<u>Source</u>	<u>K<sub>d</sub>, pM</u>	<u>Relative cell proliferation<sup>†</sup></u>
	Fab	<i>E. coli</i>	570	91
	F(ab') <sub>2</sub> <sup>‡</sup>	<i>E. coli</i>	290	53
	F(ab') <sub>2</sub>	293 cells	300	50

10

\* K<sub>d</sub> values for the p185<sup>HER2</sup> ECD were determined as previously described (5) and the standard error of the estimates are  $\leq \pm 10\%$ .

15

<sup>†</sup> Proliferation of SK-BR-3 cells incubated for 96 hr with huMAb4D5 variants shown as a percentage of the untreated control as described (5). Data represent the maximal anti-proliferative effect for each variant calculated as the mean of triplicate determinations at a fragment concentration of 10  $\mu$ g/ml. Data are all taken from the same experiment and the estimated standard error  $\leq \pm 15\%$ .

20

<sup>‡</sup> Cys Ala Ala variant.

The binding affinity of huMAb4D5-8  $F(ab')_2$  antibody fragment for  $p185^{HER2}$  ECD is identical to that of the corresponding fragment derived from limited proteolysis of whole antibody expressed in mammalian cells. The bivalent  $F(ab')_2$  antibody 5 fragment derived from *E. coli* has identical anti-proliferative activity with SK-BR-3 cells to both the intact bivalent huMAb4D5-8 parent antibody derived from 293 cells (6) and the  $F(ab')_2$  antibody fragment derived from limited pepsin digestion of intact antibody. In contrast, the monovalent Fab molecule does not significantly affect the 10 growth of SK-BR-3 cells. This suggests that the crosslinking of  $p185^{HER2}$  on the surface of cells may be required for inhibiting their proliferation. It appears exceedingly unlikely that the 2-fold weaker antigen binding affinity of the Fab compared to the  $F(ab')_2$  could account for the lack of anti-proliferative activity of the 15 Fab at concentrations up to several hundred fold above the  $K_d$ . Furthermore, it is possible to block the anti-proliferative activity of the bivalent parent antibody, muMAb4D5, with a monovalent huMAb4D4 Fab fragment.

In this example, the expression titer of functional Fab fragments was increased compared to the reports in the literature for *E. coli*, from about 1000-fold to 1 to 2 grams per liter. Additionally, Fab' molecules were recovered. This 20 enhanced expression is very likely due only in part to the higher cell densities (10 to 20-fold) and more precisely controlled environment of the fermentor than the simple shake flask, the very tight control of pre-induction expression and the characteristics of the humanized variable domain sequence employed. The titer of Fab' obtained here using the phoA promoter are surprisingly high and may result from a combination 25 of the use of this powerful promoter in a low copy number (pBR322-based) vector. Thus the repressor is not titrated to less effective levels. The cells also could be transformed to make surplus repressor. What is important is that the promoter be inactive prior to induction. It is possible that vector design and the high thermal 30 stability of the Fab fragment ( $T_m > 80^\circ\text{C}$ ) also may be important. This system should greatly facilitate clinical or biophysical studies requiring large quantities of antibody fragments.

Similar high expression levels have been observed for several additional variants of huMAb4D5-8 Fab which contain one or more amino acid changes in the antigen binding loops or nearby framework residues (6). Two 35 alternative antigen-binding specificities have been recruited into huMAb4D5-8 Fab' by judicious replacement of antigen binding residues and found to give high

expression titers. This is consistent with the notion that the framework of a polypeptide comprising substantially huMAb4D5-8 Fab' might be generically useful for highly expressed humanized Fab' molecules.

5 The huMAb4D5-8 Fab' Cys Ala Ala shows very little tendency to form  $F(ab')_2$  *in vivo* despite the apparently quantitative formation of intra-domain disulfides in the variable regions. However, *in vitro*  $F(ab')_2$  forms readily by air oxidation of Fab'-SH at pH 7.5 in the absence of EDTA at concentrations that are at least 10-fold lower than are found *in vivo*.

10 Without committing to any particular mechanism, it is believed that the redox potential of the periplasmic space of *E. coli* is sufficiently oxidizing to allow formation of the intra-domain disulfide bonds but not the inter-heavy chain disulfide which is presumably thermodynamically less favorable. Nevertheless, recovery of functional Fab' fragments secreted into the periplasmic space of *E. coli* with the unpaired hinge cysteine mainly as the free thiol provides the essential starting 15 material for directed coupling (23, 32, 33).

15 We have additionally exploited the free thiol for immobilization of the huMAb4D5-8 Fab' fragment on an activated thiol support as previously described (34), enabling the p185<sup>HER2</sup> ECD to be affinity purified from solution. The free hinge thiol has also been used for attachment of fluorescent probes for fluorescence-activated cell sorting. It is also 20 within the scope of this invention to use the free cysteinyl thiol for the site-directed attachment of radionuclides for imaging or therapy. This would offer the advantage over conventional labelling strategies of a defined stoichiometry and attachment site without the risk of compromising antigen binding affinity.

25

#### Bibliography

1. B. M. Fendly *et al.*, *Cancer Res.* **50**: 1550 (1990).
2. R. M. Hudziak *et al.*, *Molec. Cell. Biol.* **9**: 1165 (1989).
- 30 3. R. Lupu *et al.*, *Science* **249**: 1552 (1990).
5. D. J. Slamon *et al.*, *Science* **235**: 177 (1987); D. J. Slamon *et al.*, *Science* **244**: 707 (1989).
- 35 6. P. Carter *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**(10):4285-9 (1992).
7. P. T. Jones *et al.*, *Nature* **321**: 522 (1986); L. Riechmann *et al.*, *Nature* **332**: 323 (1988); M. Verhoeven *et al.*, *Science* **239**: 1534 (1988).
- 40 8. R. D. Blumenthal *et al.*, *Adv. Drug Del. Rev.* **4**: 279 (1990).

9. P. Parham in *Cellular Immunology* (E. M. Weir, Ed., Blackwell Scientific, CA) 4th edition, vol 1 chapter 14 (1983).

5 10. A. Skerra and A. Plückthun, *Science* **240**: 1038 (1988); M. Better *et al.*, *ibid* p. 1041.

10 11. C. N. Chang *et al.*, *Gene* **44**: 121 (1986).

12. R. N. Picken *et al.*, *Infect. Immun.* **42**: 269 (1983).

10 13. W. Palm and N. Hilschmann, *Z. Physiol. Chem.* **356**: 167 (1975).

14. J. W. Ellison *et al.*, *Nucleic Acids Res.* **10**: 4071 (1982).

15 15. F. Bolivar *et al.*, *Gene* **2**: 95 (1977).

16. S. Scholtissek, *et al.*, *Nucleic Acids Res.* **15**: 3185 (1987)

17. E. A. Kabat, *et al.*, *Sequences of Proteins of Immunological Interest* 3rd edition (National Institutes of Health, Bethesda, MD, 1987).

20 18. J. A. Wells *et al.*, *Gene* **34**: 315 (1985).

19. R.A. Maurer, Ph.D. thesis, Harvard University (1978).

25 20. The huMAb4D5-8 Fab' fragment was expressed in *E. coli* strain 25F2 derived from the strain RV308 (ref. 22, ATCC#31608) by inactivating the *tonA* gene. Cells were grown for 32-40 hours in an aerated 10 liter fermentor at 37°C at a low agitation rate (650 rpm, *Kla* ~600 mmol  $1^{-1}$   $hr^{-1}$   $atm^{-1}$ ) in a medium that initially contained 12 g  $1^{-1}$  digested casein, 17 mM glucose, 2.4 mM isoleucine hydrochloride, 47 mM  $(NH_4)_2SO_4$ , 10 mM  $NaH_2PO_4$ , 18 mM  $K_2HPO_4$ , 4.1 mM trisodium citrate, 12 mM  $MgSO_4$ , 125  $\mu$ M  $FeCl_3$  and 20  $\mu$ M each of  $ZnSO_4$ ,  $MnSO_4$ ,  $CuSO_4$ ,  $CoCl_2$ ,  $H_3BO_3$  and  $NaMoO_4$ ) plus 12 mg  $1^{-1}$  tetracycline and received automated feeds of ammonia to maintain the pH at 7.0 and also glucose to maintain a slight excess or avoid anaerobiosis depending upon cell density of 80 to 100  $OD_{550}$ . The cell density at harvest is usually 120 to 150  $OD_{550}$ .

30 35 40 21. G. L. Ellman, *Arch. Biochem. Biophys.* **82**: 70 (1959).

23. M. Brennan *et al.*, *Science* **229**: 81 (1985).

24. E. Lamoyi, *et al.*, *Methods Enzymol.* **121**: 652 (1986).

45 25. T. E. Creighton, *Protein Structure, a Practical Approach* (IRL Press, Oxford, UK, 1990), p. 157.

26. P. Matsudaira, *J. Biol. Chem.* **262**: 10035 (1987)

50 27. S. Moore and W.H. Stein, *Methods Enzymol.* **6**: 819 (1963)

32. M. J. Glennie *et al.*, *J. Immunol.* **139**: 2367 (1987); M. J. Glennie *et al.*, *J.*

-37-

*Immunol.* **141**: 3662 (1988).

33. T. Nitta *et al.*, *J. Immunol.* **19**: 1437 (1989). T. Nitta *et al.*, *Lancet* **335**: 368 (1990).

5 34. P. Carter and J. A. Wells, *Science* **237**: 394 (1987).

-38-

I claim:

1. A method comprising expressing and secreting into the periplasm of a recombinant microbial cell culture a Fv polypeptide containing an immunoglobulin heavy chain Fv region and an immunoglobulin light chain Fv region, said light or heavy chain also comprising an unpaired cysteinyl residue as a free thiol, and recovering said polypeptide under conditions that substantially maintain said cysteinyl residue as the free thiol.
- 10 2. A method for the production of a Fab' antibody polypeptide having at least one hinge region cysteine present as a free thiol (Fab'-SH), comprising the steps of:
  - 15 a. expressing nucleic acid encoding an immunoglobulin presequence comprising Fab' in a microbial host cell culture transformed with a vector comprising said nucleic acid operably linked to control sequences recognized by the host cell transformed with the vector, under conditions suitable for the secretion of Fab' to the periplasmic space of the host cell and formation of Fab'-SH; and
  - 20 b. recovering Fab'-SH from said host cell.
3. The method of claim 2 wherein said Fab'-SH is not exposed to reducing conditions during step b.
- 25 4. The method of claim 2 wherein said Fab'-SH has only one hinge region cysteine.
5. The method of claim 4 wherein said Fab'-SH is recovered under 30 conditions suitable for maintaining the hinge cysteinyl thiol in protonated form.
6. The method of claim 2 wherein a metal ion chelating agent is present during the culturing of the transformed cell.

-39-

7. The method of claim 2 wherein a metal ion chelating agent is present during the recovery of said Fab'-SH.
- 5 8. The method of claim 2 wherein a protease inhibitor is present during the recovery of said Fab'-SH.
9. The method of claim 8 wherein the protease inhibitor is present during the culturing of the transformed cell.
- 10 10. The method of claim 2, wherein said Fab'-SH is recovered by freeze-thawing the host cell and subjecting it to osmotic shock in the presence of lysozyme.
- 15 11. The method of claim 2 wherein Fab' comprises the C-terminal amino acid sequence Cys Ala Ala.
12. The method of claim 2 wherein said microbial host is *E. coli*.
13. A method for the preparation of a polypeptide comprising F(ab')<sub>2</sub> comprising the steps of:
  - 20 a. expressing nucleic acid encoding a first a Fab' in a microbial host cell culture transformed with a vector comprising the nucleic acid operably linked to control sequences recognized by the host cell transformed with the vector, under conditions suitable for the secretion of Fab' to the periplasmic space of the host cell and formation of Fab'-SH, said first Fab' being capable of binding a first epitope;
  - 25 b. expressing nucleic acid encoding a second a Fab' in a microbial host cell culture transformed with a vector comprising the nucleic acid operably linked to control sequences recognized by the host cell transformed with the vector, under conditions suitable for the secretion of Fab' to the periplasmic space of the host cell and formation of Fab'-SH, said second Fab' being
- 30
- 35

-40-

capable of binding a second epitope;

5           c.     recovering said first and second Fab'-SH from said host cells; and

          d.     forming a covalent bond between the cysteinyl free thiols of said first and second Fab'-SH to form bivalent F(ab')<sub>2</sub>.

10           14.    The method of claim 12, wherein said covalent bond is a disulfide bond.

15           15.    The method of claim 13 wherein the bond formation between the first and second Fab'-SH comprises the following steps:

          a.     reacting the first Fab'-SH with (i) 5,5'-dithiobis (2-nitrobenzoic acid) (DTNB) to form a thionitrobenzoate derivative Fab'-TNB or (ii) a bifunctional maleimide;

          b.     directly coupling said Fab'-TNB or maleimidated Fab' to the second Fab'-SH to form a F(ab')<sub>2</sub>; and

          c.     recovering said F(ab')<sub>2</sub>.

20           16.    The method of claim 12, wherein the epitopes bound by the two Fab' are located on the same antigen.

25           17.    The method of claim 12, wherein the epitopes bound by the two Fab' are located on different antigens.

          18.    The method of claim 12 wherein each Fab' is present in the culture at levels in excess of about 1 gram/liter.

30           19.    The method of claim 12 wherein each Fab' comprises complementarity determining region amino acid sequence from a non-human immunoglobulin and other amino acid sequences from a human immunoglobulin.

          20.    The method of claim 12 wherein the amino acid sequences are obtained from IgG.

-41-

21. A composition comprising  $F(ab')_2$  which is

- a. essentially free of  $F(ab')_2$  Fv regions containing cysteinyl residues with derivatized sulphydryl groups, except for any disulfide bond found in the native Fv regions,
- b. entirely free of  $F(ab')_2$  having hinge region intrachain disulfide bonds,
- c. entirely free of contaminating arsenite, and
- d. entirely homogenous as to the heavy chain C-terminal amino acid residue.

10

22. A method for high yield production of an immunoglobulin polypeptide comprising culturing a host cell transformed with nucleic acid encoding an immunoglobulin polypeptide under the transcriptional control of an inducible promoter/operator system whereby expression of the polypeptide is repressed prior to induction sufficient to permit post-induction polypeptide levels in the cell culture of greater than about 1 gram of polypeptide per liter of cell culture.

15

20

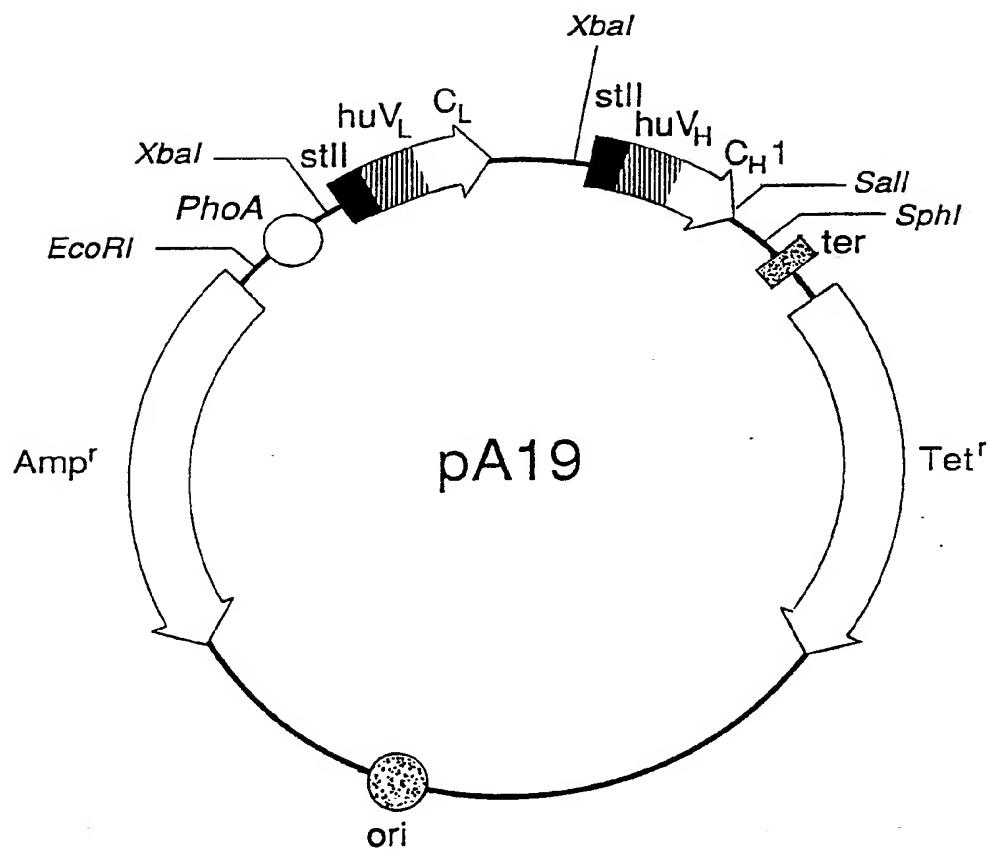
23. The method of claim 2 wherein said Fab'-SH has more than one hinge region cysteine.

25

24. The method of claim 2 wherein the light and heavy chains of said Fab' are not covalently bound.

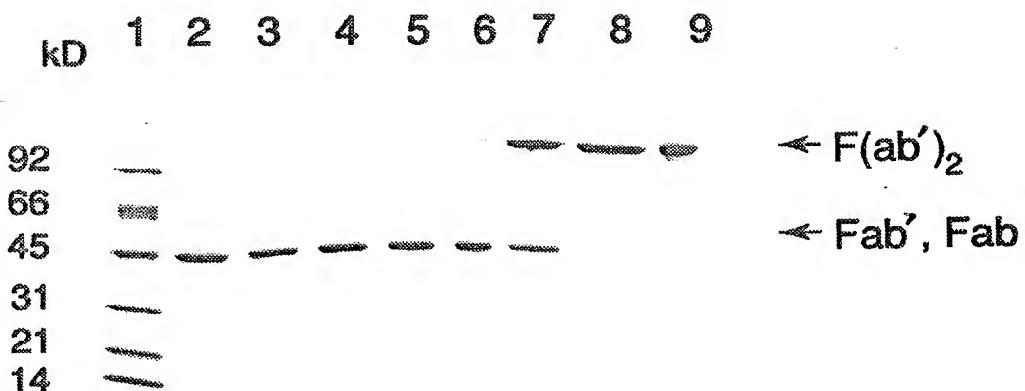
1/2

FIG. 1



2/2

## FIG. 2



## INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 92/07986

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)<sup>6</sup>

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.C1. 5 C12N15/13; C12P21/08

## II. FIELDS SEARCHED

Minimum Documentation Searched<sup>7</sup>

Classification System	Classification Symbols		
Int.C1. 5	C07K	;	C12N ; C12P

Documentation Searched other than Minimum Documentation  
to the Extent that such Documents are Included in the Fields Searched<sup>8</sup>III. DOCUMENTS CONSIDERED TO BE RELEVANT<sup>9</sup>

Category <sup>10</sup>	Citation of Document, <sup>11</sup> with indication, where appropriate, of the relevant passages <sup>12</sup>	Relevant to Claim No. <sup>13</sup>
P, X	BIOTECHNOLOGY vol. 10, no. 2, February 1992, NEW YORK US pages 163 - 167 CARTER, PAUL; KELLEY, ROBERT F.; RODRIGUES, MARIA L.; SNEDECOR, BRAD; COVARRUBIAS, MICHAEL; VELLIGAN, MARK D.; WONG, WAI LEE T.; R 'High level Escherichia coli expression and production of a bivalent humanized antibody fragment' see the whole document ---	1-24
X	BIOCHEMISTRY. vol. 29, 1990, EASTON, PA US pages 1362 - 1367 GLOKSHUBER, R. ET AL. 'A comparison of strategies to stabilize Immunoglobulin Fv-Fragments.' see the whole document ---	1,22 -/-

<sup>10</sup> Special categories of cited documents :<sup>10</sup>

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

- "&" document member of the same patent family

## IV. CERTIFICATION

Date of the Actual Completion of the International Search

11 JANUARY 1993

Date of Mailing of this International Search Report

29.01.93

International Searching Authority

EUROPEAN PATENT OFFICE

Signature of Authorized Officer

S.A. NAUCHE

III. DOCUMENTS CONSIDERED TO BE RELEVANT		(CONTINUED FROM THE SECOND SHEET)
Category <sup>a</sup>	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
P, X	<p>PROTEIN ENGINEERING  vol. 4, no. 8, December 1991,  pages 971 - 979  SKERRA, ARNE; PLÜCKTHUN, ANDREAS  'Secretion and in vivo folding of the Fab  fragment of the antibody McPC603 in  Esherichia Coli : Influence of disulphides  and cis-prolines.'  see the whole document</p> <p>---</p>	1
X	<p>BEHRING INST. MITT.  vol. 78, 1985,  pages 118 - 132  PAULUS, H. ET AL. 'Preparation and  biomedical applications of bispecific  antibodies.'  see the whole document</p> <p>---</p>	21
X	<p>METHODS IN ENZYMOLOGY  vol. 178, 1989, ACADEMIC PRESS, US  pages 497 - 515  PLÜCKTUHN, A.; SKERRA, A. 'Expression of  functional Antibody Fv and Fab fragments  in Esherichia Coli.'  cited in the application  See The whole document, especially figure  2</p> <p>-----</p>	1, 22